

**IN THE UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF DELAWARE**

THOMSON REUTERS ENTERPRISE)	
CENTRE GMBH and WEST PUBLISHING)	
CORPORATION,)	C.A. No. 20-613-SB
)	
Plaintiffs/Counterdefendants,)	
)	
v.)	PUBLIC VERSION
)	
ROSS INTELLIGENCE INC.,)	
)	
Defendants/Counterclaimant.)	

**ROSS INTELLIGENCE INC.'S OPENING BRIEF IN SUPPORT OF ITS MOTION FOR
SUMMARY JUDGMENT ON ITS AFFIRMATIVE DEFENSE OF FAIR USE**

OF COUNSEL:

Warrington S. Parker III
Joachim B. Steinberg
Jacob Canter
CROWELL & MORING LLP
3 Embarcadero Ctr., 26th Floor
San Francisco, CA 94111
Tel: (415) 986-2800

Keith J. Harrison
Crinesha B. Berry
CROWELL & MORING LLP
1001 Pennsylvania Avenue NW
Washington, DC 20004
Tel: (202) 624-2500

David E. Moore (#3983)
Bindu A. Palapura (#5370)
Andrew L. Brown (#6766)
POTTER ANDERSON & CORROON LLP
Hercules Plaza, 6th Floor
1313 N. Market Street
Wilmington, DE 19801
Tel: (302) 984-6000
dmoore@potteranderson.com
bpalapura@potteranderson.com
abrown@potteranderson.com

*Attorneys for Defendant/Counterclaimant ROSS
Intelligence, Inc.*

Dated: October 1, 2024
11790294 / 20516.00001
Public Version Dated: October 9, 2024

TABLE OF CONTENTS

	Page
INTRODUCTION AND SUMMARY OF ARGUMENT	1
NATURE AND STAGE OF PROCEEDINGS	3
STATEMENT OF FACTS	3
A. ROSS Wanted to Make Legal Services Cheaper and More Accessible	3
B. What Plaintiffs State in Discovery Responses that ROSS Has Infringed.....	6
1. Plaintiffs State that ROSS Directly Infringed the Selection and Arrangement of Headnotes and Case Passages that were Copied into the Bulk Memos.....	6
2. Plaintiffs State that ROSS Directly Infringed the Text of Headnotes Copied into the Bulk Memos	7
3. Plaintiffs State that ROSS Directly Infringed the Selection and Arrangement of Headnotes, Case Passages, and Cases as Part of the Key Number System that were Copied into the Bulk Memos	7
4. Plaintiffs State that ROSS Directly Infringed the Key Number System Topics that were Listed in the Bulk Memo File Names	8
5. Plaintiffs State that ROSS Directly Infringed the Copyrighted Content that was within the Cases Given to ROSS as Part of the Classifier Project	8
6. Plaintiffs State that ROSS Indirectly Infringed Due to LegalEase’s Direct Infringement of Copyrighted Content that LegalEase Accessed on Westlaw.....	8
7. Plaintiffs Now Also State that ROSS Infringed “Westlaw”	9
C. The Westlaw Platform and Its History.....	9
1. Judicial Opinions are Available on Westlaw	10
2. Headnotes are Available on Westlaw	11
3. The Key Number System is Available on Westlaw.....	11
4. Other “Editorial Enhancements” are Available on Westlaw	12
D. Plaintiffs’ Copyright in Westlaw Content.....	12
E. The Search Technology Employed by Plaintiffs on Westlaw	13
1. WestSearch – Algorithms Trained on User Behavior, Citation Networks, and Key Numbers; Searches All Legal Content; Results are Documents	13
2. WestSearch Plus – Trained with Headnotes as the Answers in QA Pairs; Only Responds to Questions that Plaintiffs Pose; Results are Headnotes	15

F.	To Develop Its Legal Research Tool, ROSS Needed Pairs of Legal Questions and Judicial Quotes that Answer those Legal Questions; It Got This in the Form of Memoranda	16
G.	LegalEase Prepared the Memos Containing the Question-Answer Pairs	17
H.	How LegalEase Prepared the Memos and What LegalEase Gave ROSS	18
I.	To Complete the Memos, LegalEase Drafted the Legal Questions to Be as Close as Possible to the Language in Judicial Opinions and Took Verbatim Paragraphs from Judicial Opinions for the Answers	19
J.	The Completed Memos that LegalEase Gave to ROSS Did Not Contain Plaintiffs' Copyrighted Content	20
K.	The Additional Items Provided to ROSS by LegalEase that ROSS Rejected	20
1.	List of 91 Legal Topics	20
2.	500 Judicial Opinions	21
L.	The Development of ROSS's Legal Research Tool and How it Operates	21
1.	ROSS First Extracted the Question-Answer Pairs from the Memos, Discarded Some of the Pairs, and Randomly Separated the Pairs for Different Uses	22
2.	ROSS Then Prepared the Question-Answer Pairs for Computational Analysis of the Text's Non-Expressive Characteristics.....	23
3.	ROSS Featurizes the Question-Answer Pairs—Extracting Only Numerical Representations of the Relationships Among the Language and Discarding the Actual Text.....	23
4.	ROSS Uses the Numbers that Represent Abstract Characteristics of the Text to “Train” the Machine-Learning Model.....	25
5.	How ROSS's Search Tool Operates When a User Makes a Query – Users Can Ask Any Questions; the Outputs are Only Judicial Quotes	26
	ARGUMENT	27
I.	FACTOR 2: NATURE OF THE COPYRIGHTED WORK	28
II.	FACTOR 3: AMOUNT AND SUBSTANTIALITY OF THE WORK TAKEN.....	31
III.	FACTOR 1: THE PURPOSE AND CHARACTER OF THE USE.....	33
IV.	FACTOR 4: THE EFFECT OF THE USE ON THE MARKET	36
	CONCLUSION.....	40

TABLE OF AUTHORITIES

	Page(s)
Cases	
<i>ACT, Inc. v. Worldwide Interactive Network, Inc.</i> , 46 F.4th 489 (6th Cir. 2022)	35
<i>Andy Warhol Found. for Visual Arts, Inc. v. Goldsmith</i> , 11 F.4th 26 (2d Cir. 2021)	34
<i>Andy Warhol Found. for the Visual Arts, Inc. v. Goldsmith</i> , 992 F.3d 99 (2d Cir. 2021).....	34
<i>Apple Barrel Prods., Inc. v. Beard</i> , 730 F.2d 384 (5th Cir. 1984)	30
<i>Authors Guild, Inc. v. HathiTrust</i> , 755 F.3d 87 (2d Cir. 2014).....	33, 36
<i>Authors Guild v. Google, Inc.</i> , 804 F.3d 202 (2d Cir. 2015).....	27, 32, 33, 36
<i>Baker v. Selden</i> , 101 U.S. 99 (1879).....	35
<i>BellSouth Advert. & Pub. Corp. v. Donnelley Info. Pub., Inc.</i> , 999 F.2d 1436 (11th Cir. 1993)	30
<i>Bikram’s Yoga Coll. of India v. Evolation Yoga, Ltd. Liab. Co.</i> , 803 F.3d 1032 (9th Cir. 2015)	35
<i>Bill Graham Archives v. Dorling Kindersley Ltd.</i> , 448 F.3d 605 (2d Cir. 2006).....	34
<i>Blanch v. Koons</i> , 467 F.3d 244 (2d Cir. 2006).....	34
<i>Brown v. Netflix, Inc.</i> , 462 F. Supp. 3d 453 (S.D.N.Y. 2020), <i>aff’d</i> , 855 F. App’x 61 (2d Cir. 2021)	37
<i>Cambridge Univ. Press v. Patton</i> , 769 F.3d 1232 (11th Cir. 2014)	38
<i>Campbell v. Acuff-Rose Music, Inc.</i> , 510 U.S. 569 (1994).....	27, 28, 33

<i>Cariou v. Prince</i> , 714 F.3d 694 (2d Cir. 2013).....	34
<i>Craft Smith, LLC v. EC Design, LLC</i> , 969 F.3d 1092 (10th Cir. 2020)	30
<i>Dun & Bradstreet Software Servs., Inc. v. Grace Consulting, Inc.</i> , 307 F.3d 197 (3d Cir. 2002).....	29
<i>Experian Info. Sols., Inc. v. Nationwide Mktg. Servs. Inc.</i> , 893 F.3d 1176 (9th Cir. 2018)	30
<i>Georgia v. Public.Resource.Org, Inc.</i> , 140 S. Ct. 1498 (2020).....	40
<i>Google LLC v. Oracle Am., Inc.</i> , 593 U.S. 1 (2021).....	29, 35, 39
<i>Harper & Row Publishers, Inc. v. Nation Enterprises</i> , 471 U.S. 539 (1985).....	30
<i>Hustler Mag. Inc. v. Moral Majority Inc.</i> , 796 F.2d 1148 (9th Cir. 1986)	28
<i>Intervest Constr., Inc. v. Canterbury Est. Homes, Inc.</i> , 554 F.3d 914 (11th Cir. 2008)	28
<i>Keck v. Mix Creative Learning Ctr., L.L.C.</i> , No. 23-20188, --- F.4th ---, 2024 WL 4222188 (5th Cir. Sept. 18, 2024).....	38
<i>Kelly v. Arriba Soft Corp.</i> , 336 F.3d 811 (9th Cir. 2003)	34
<i>Matthew Bender & Co. v. W. Publ’g Co.</i> , 158 F.3d 674 (2d Cir. 1998).....	28, 32
<i>Matthew Bender & Co. v. W. Publ’g Co.</i> , 158 F.3d 693 (2d Cir. 1998).....	2, 30
<i>Miller v. Universal City Studios, Inc.</i> , 650 F.2d 1365 (5th Cir. 1981)	30
<i>Mitel, Inc. v. Iqtel, Inc.</i> , 124 F.3d 1366 (10th Cir. 1997)	28
<i>Monsarrat v. Newman</i> , 28 F.4th 314 (1st Cir. 2022).....	29

<i>New Era Publ'ns Int'l, ApS v. Carol Publ'g Grp.</i> , 904 F.2d 152 (2d Cir. 1990).....	32
<i>Nihon Keizai Shimbun, Inc. v. Comline Bus. Data, Inc.</i> , 166 F.3d 65 (2d Cir. 1999).....	31
<i>S. Credentialing Support Servs., L.L.C. v. Hammond Surgical Hosp., L.L.C.</i> , 946 F.3d 780 (5th Cir. 2020)	30
<i>Sega Enters. v. Accolade, Inc.</i> , 977 F.2d 1510 (9th Cir. 1992)	34, 36, 37, 39
<i>Seltzer v. Green Day, Inc.</i> , 725 F.3d 1170 (9th Cir. 2013)	34
<i>SOFA Ent., Inc. v. Dodger Prods., Inc.</i> , 709 F.3d 1273 (9th Cir. 2013)	29
<i>Sony Comput. Entm't, Inc. v. Connectix Corp.</i> , 203 F.3d 596 (9th Cir. 2000)	<i>passim</i>
<i>Swatch Grp. Mgmt. Servs. Ltd. v. Bloomberg L.P.</i> , 756 F.3d 73 (2d Cir. 2014).....	29, 38
<i>Taylor v. Commissioner</i> , 51 F.2d 915 (3d Cir. 1931).....	35
<i>TCA Television Corp. v. McCollum</i> , 839 F.3d 168 (2d Cir. 2016).....	38
<i>A.V. ex rel. Vanderhye v. iParadigms, LLC</i> , 562 F.3d 630 (4th Cir. 2009)	33
<i>Whiddon v. Buzzfeed, Inc.</i> , 638 F. Supp. 3d 342 (S.D.N.Y. 2022).....	31, 32
<i>White v. W. Publ'g Corp.</i> , 29 F. Supp. 3d 396, 399 (S.D.N.Y. 2014)	33
Statutes	
17 U.S.C. § 102.....	28
17 U.S.C. § 107.....	27
Other Authorities	
ADBI AIDID & BENJAMIN ALARIE, THE LEGAL SINGULARITY: HOW ARTIFICIAL INTELLIGENCE CAN MAKE LAW RADICALLY BETTER 54-55 (2023)	4

Copyright Act.....	2, 35
4 Melville B. Nimmer & David Nimmer, NIMMER ON COPYRIGHT § 13F.06[A] (rev. ed. 2024)	<i>passim</i>
Olufunmilayo B. Arewa, <i>Open Access in A Closed Universe: Lexis, Westlaw, Law Schools, and the Legal Information Market</i> , 10 Lewis & Clark L. Rev. 797, 832 (2006)	4
Thomson Reuters	10, 13, 14, 18

INTRODUCTION AND SUMMARY OF ARGUMENT

If ROSS did infringe Plaintiffs' copyrights, ROSS's use constituted fair use. The factors considered weigh in ROSS's favor.

Factor One: *Purpose and Character of the Use.* ROSS set out to reduce the cost of legal research and increase the accessibility of the law by adapting artificial intelligence ("AI") technology to legal search engines. ROSS enlisted LegalEase to use what Plaintiffs claim are copyrighted materials for their intended use: the location of judicial opinions for the purpose of creating memos that contained legal questions and quotes from judicial opinions. This copying was an intermediate step to enable the memos to teach the ROSS legal search engine.

To teach the legal search engine, ROSS took the memos and used [REDACTED] different mathematical relationship models that allowed the creation of software code that reflects how words relate one to another mathematically. The memos were not digitized—reduced to ones and zeros. Once the legal search engine learned how words relate to each other mathematically, the memos were discarded. They do not appear in the legal search engine. They cannot be found in the software code. There can be no dispute over this because Plaintiffs acknowledge the same: after the AI legal search tool is [REDACTED] it is [REDACTED] and the question-answer pairs are no longer used or relevant.¹

When a user ran a search, it used the code that only knew the mathematical relations of words across judicial opinions obtained from third parties that were devoid of any Westlaw Content. For these reasons, no headnotes or any other Westlaw Content could be returned by the search.

¹ See Declaration of Jacob Canter ("Canter Decl.") Ex. 1 (I. Moulinier Dep. Tr.), 61:22-62:6.

Factor Two: *Nature of the Copyrighted Work.* This focuses on how creative a work is. The less creative the more this factor leans in favor of fair use. The Westlaw Content is a method of organizing cases so that lawyers can locate them. The Westlaw Content is registered as a compilation. While protected as a compilation as a whole by the Copyright Act, the scope of protection afforded is thin. *Matthew Bender & Co. v. W. Publ'g Co.*, 158 F.3d 693, 707 (2d Cir. 1998) (“West has a thin copyright in its compilations.”). The Westlaw Content is constrained by Plaintiffs’ choice to reflect the words of judicial opinions so that lawyers can locate the judicial opinions and their contents. Plaintiffs acknowledge that its Westlaw Content is constrained by the language of the law. Even were headnotes treated individually, they are less creative as they too are constrained by the words of judicial opinions.

Factor Three: *Amount and Substantiality of The Portion Used.* This factor asks how much of the copyrighted material as compared to the whole was used. This factor too weighs in ROSS’s favor. As of [REDACTED], there were [REDACTED] unique headnotes in Westlaw. Plaintiffs claim that LegalEase infringed [REDACTED] headnotes in creating memos. Plaintiffs identify [REDACTED] subtopics and key numbers copied to create those memos. Only considering the headnotes, this is only 0.07597% of all headnotes. Considering only the subtopics and key numbers, the [REDACTED] traversed during the memo creation process is only 5.47% of the [REDACTED] topics, subtopics, and key numbers in Westlaw.

These minuscule numbers do not even consider the items within the “works as a whole,” which is the relevant inquiry. The Westlaw compilation contains 11,205,374,706 data records. Using that as the denominator, what is infringed constitutes .000444% of what Plaintiffs have registered.

Factor Four: *Effect of the Use on the Market for the Work.* This factor asks whether ROSS's use had an effect on the market for or value of the work. ROSS's legal search engine could not impact the market in Westlaw Content because ROSS's search engine neither employs them nor returns them as search results. Those wishing to use Westlaw Content have only one place to go: Westlaw.

Plaintiffs effectively concede that ROSS does not injure their market in the copyrighted materials as they appear in their legal search engine because they point to a different one: the AI training data market. However, ROSS was never going to sell AI training data and did not. There is no evidence that Plaintiffs ever did or ever will sell the copyrighted materials at issue here as AI training data. That Plaintiffs use their copyrighted materials as training data is irrelevant as Plaintiffs have not been deprived of that opportunity right now or in the future. Lastly, Plaintiffs' focus on the AI training data is misplaced. This factor looks at the final result of the use of the copyrighted materials. The final result was not AI training data. The final result was a legal search engine.

For these reasons, summary judgment should be granted.

NATURE AND STAGE OF PROCEEDINGS

On August 29, 2024, the Court granted the Parties' request that opening summary judgment briefs not exceed 40 pages and be filed by no later than October 1, 2024. *See* D.I. 667.

STATEMENT OF FACTS²

A. ROSS Wanted to Make Legal Services Cheaper and More Accessible

² Declarations of Jacob Canter ("Canter Decl."), Jimoh Ovbiagele ("Ovbiagele Decl."), Joseph Marks ("Marks Decl."), Richard Leiter ("Leiter Decl."), Barbara Frederiksen-Cross ("Frederiksen-Cross Decl."), and Alan Cox ("Cox Decl.") have been prepared and filed in support of every exhibit in this brief.

Jimoh Ovbiagele, one of ROSS's co-founders and the creator of ROSS's technology, has personal experience with the need to make legal services cheaper and more accessible. Canter Decl. Ex. 2 (J. Ovbiagele Dep. Tr.), 38:17-25 ("[REDACTED] [REDACTED]"); Ovbiagele Decl. Ex. 3 (ROSS-010129852), at -852 (need for cheaper legal research). For this reason, he and co-founder Andrew Arruda started ROSS in 2014. Canter Decl. Ex. 4 (A. Arruda Dep. Tr.), 11:21-12:4; Ovbiagele Decl. Ex. 5 (ROSS-001397429), at -431 [REDACTED] [REDACTED] [REDACTED]).

To make legal services cheaper and more accessible, ROSS sought to improve how lawyers conduct legal research. *See, e.g.*, Ovbiagele Decl. Ex. 6 (ROSS-023175123), at -123 (“ROSS is an Artificial Intelligence (AI) system designed to improve the efficiency, accuracy and profitability of legal research.”); *id.* Ex. 3 (ROSS-010129852), at -852 (Mr. Ovbiagele “think[s] ROSS] will allow legal services to be delivered at a more affordable price.”); *id.* Ex. 5 (ROSS-001397429), at -431; Canter Decl. Ex. 2 (J. Ovbiagele Dep. Tr.), 38:2-6.

That legal research is expensive and difficult is broadly acknowledged. *See, e.g.*, ADBI AIDID & BENJAMIN ALARIE, THE LEGAL SINGULARITY: HOW ARTIFICIAL INTELLIGENCE CAN MAKE LAW RADICALLY BETTER 54-55 (2023); Olufunmilayo B. Arewa, *Open Access in A Closed Universe: Lexis, Westlaw, Law Schools, and the Legal Information Market*, 10 Lewis & Clark L. Rev. 797, 832 (2006). This expense and difficulty is made worse by methods for legal research that rely on innovations from the 19th Century. Canter Decl. Ex. 7 (TR-0000512), at -

[REDACTED] *id.* Ex. 8 (TRCC-00686613), at -625 ([REDACTED])

[REDACTED] *id.* at -627 ([REDACTED])

[REDACTED] *id.* Ex. 9 (TRCC-00608822), at -

824.

ROSS planned to change how people search for cases by allowing its platform users to ask natural language queries where the responses are the paragraphs of judicial opinions that directly respond to the question. Ovbiagele Decl. Ex. 10 (ROSS-003487253), at -290; Ovbiagele Decl. Ex. 11 (ROSS-023179081), at -087; *id.* Ex. 12 (ROSS-003515560), at -560. This was the vision: No more Boolean searches or keywords. *Id.* Ex. 13 (ROSS-009542437), at -437 (“[REDACTED]

[REDACTED] (emphasis in original). No more struggling through a judicial opinion to discover the correct passage. *Id.* Ex. 5 (ROSS-001397429), at -429.

By making research less time-consuming for lawyers, the price for legal services would drop. Ovbiagele Decl. Ex. 5 (ROSS-001397429), at -431 (ROSS “could make possible hiring a lawyer in the first place - currently out of the reach of many people.”); *id.* Ex. 3 (ROSS-010129852), at -852.

To accomplish these goals, ROSS developed what was then-groundbreaking technology: a search tool for the legal research domain built upon deep learning techniques that use a machine-learning model. Canter Decl. Ex. 2 (J. Ovbiagele Dep. Tr.), 41:5-13. “[REDACTED]

[REDACTED] *Id.* 41:20:42:1.

ROSS did not want human intermediated content to have any impact on how its tool operated. *See* Ovbiagele Decl. ¶ 5 (“ROSS’s legal search engine took the language of judicial opinions and turned them into statistical measurements for the purpose of creating a legal search engine free of human intermediated content.”). Rather, ROSS built the tool so that the search results were driven entirely by a machine-based analysis of the judicial text. *See id.* ¶¶ 4-5, 23-29 (describing ROSS’s search engine); Marks Decl. ¶¶ 8-9, 13-24, 29-36.

B. What Plaintiffs State in Discovery Responses that ROSS Has Infringed³

1. Plaintiffs State that ROSS Directly Infringed the Selection and Arrangement of Headnotes and Case Passages that were Copied into the Bulk Memos

Plaintiffs state that ROSS directly infringed the [REDACTED]

[REDACTED] Ovbiagele Decl. Ex. 14 (Third Suppl. Resp. to Interrog. 1), at 70. Plaintiffs state that the selection and arrangement of headnotes and corresponding case passages in a judicial opinion is copyrightable because [REDACTED] were made to decide what parts of a case to make into a headnote, what the headnote should say, and where the headnote should be linked in the opinion. *See id.* at 30. Every headnote and corresponding case passage that ROSS allegedly infringed is identified in the most recent expert report served by Dr. Jonathan Krein. *Id.*

³ [REDACTED] Order, July 25, 2022, D.I. No. 201, at 1. Plaintiffs provided a response pursuant to this order. Canter Decl. Ex. 14 (Third Suppl. Resp. to Interrog. 1), at 69-71; *see also id.* Ex. 15 (Suppl. Krein Report (Aug. 11, 2024)); *id.* Ex. 16 (Appendix B to Suppl. Krein Report).

Ex. 16 (Appendix. B to Suppl. Krein Report). The report lists [REDACTED] headnote-to-case-passage relationships. *Id.* at Column [REDACTED]

2. Plaintiffs State that ROSS Directly Infringed the Text of Headnotes Copied into the Bulk Memos

Plaintiffs state that ROSS directly infringed [REDACTED]
[REDACTED] *Id.* Ex. 14 (Third Suppl. Resp. to Interrog. 1), at 70. Dr. Krein opines that each of these headnotes are independently creative. *See id.* Ex. 15 (Suppl. Krein Report (August 11, 2024)) ¶¶ 24-27. The [REDACTED] headnotes that Plaintiffs state are independently copyrightable and that Plaintiffs state ROSS infringed are identified by Dr. Krein. *Id.* Ex. 16 (Appendix B to Suppl. Krein Report) Column [REDACTED]

3. Plaintiffs State that ROSS Directly Infringed the Selection and Arrangement of Headnotes, Case Passages, and Cases as Part of the Key Number System that were Copied into the Bulk Memos

Plaintiffs state that [REDACTED]
[REDACTED] Ovbiagele Decl. Ex. 14 (Third Suppl. Resp. to Interrog. 1), at 70. Plaintiffs state that the creative choices made here include deciding what topics or subtopics should be associated to a headnote, and deciding whether new topics or subtopics should be developed. *See id.* at 30-31. Plaintiffs state that the portions of the key number system where these headnotes, case passages, and cases are located is identified in a table that includes [REDACTED]. *Id.* at 70 n.10 (referring to table at pages 41-43 of the same interrogatory response). Dr. Krein's report identifies the location in the key number system where each headnote and associated case passage is placed. *Id.* Ex. 16 (Appendix B to Suppl. Krien Report) Column [REDACTED] In this appendix, there is significant overlap across the [REDACTED], and there are only [REDACTED] distinct topic-plus-key number locations in the key number system listed. *See* Frederiksen-Cross Decl. ¶¶ 23-24.

4. Plaintiffs State that ROSS Directly Infringed the Key Number System Topics that were Listed in the Bulk Memo File Names

Plaintiffs state that ROSS copied [REDACTED]

[REDACTED] Ovbiagele Decl. Ex. 14 (Third Suppl. Resp. to Interrog. 1), at 70.

The filename for each of the bulk memos (to the extent there was one) is listed by Dr. Krein. *Id.* Ex. 16 (Appendix B to Suppl. Krein Report) Column [REDACTED]

Many memo filenames do not contain any text that can be construed as a legal topic or West classification. *Id.*; *see also* Ovbiagele Decl. ¶ 21; Frederiksen-Cross Decl. ¶ 27. To the extent a legal topic is included in a memo filename, it is a generic legal topic consistent with the practice of law. In those cases, [REDACTED] memos contain a legal topic as a file name from topics produced by LegalEase. Frederiksen-Cross Decl. ¶ 21.

5. Plaintiffs State that ROSS Directly Infringed the Copyrighted Content that was within the Cases Given to ROSS as Part of the Classifier Project

Plaintiffs state that ROSS copied its [REDACTED]

[REDACTED] Ovbiagele Decl. *Id.* Ex. 14 (Third Suppl. Resp. to Interrog. 1), at 71. These were [REDACTED] that LegalEase provided to ROSS via email. *Id.* at 44.

6. Plaintiffs State that ROSS Indirectly Infringed Due to LegalEase's Direct Infringement of Copyrighted Content that LegalEase Accessed on Westlaw

Plaintiffs state that after LegalEase [REDACTED] to use the Westlaw platform, its [REDACTED] was infringing. Ovbiagele Decl. Ex. 14 (Third Suppl. Resp. to Interrog. 1), at 68. The infringing activity on the website includes [REDACTED]

[REDACTED] *Id.* Plaintiffs state that

[REDACTED] assisting ROSS with the classifier project, LegalEase's infringing activity was

[REDACTED]

[REDACTED]

[REDACTED] *Id.* at 68-69. Plaintiffs state that with respect to creating the bulk memos for ROSS, [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED] *Id.* at 70-71. Except for the [REDACTED] headnotes and the 5,473 key numbers that Plaintiffs claim ROSS infringed, Plaintiffs did not identify the headnotes or key numbers at issue with regard to its direct infringement theory against LegalEase or quantify those numbers. *Id.* at 68-71.

7. Plaintiffs Now Also State that ROSS Infringed “Westlaw”

After the close of discovery, Plaintiffs now allege that ROSS has directly and indirectly infringed “Westlaw.” D.I. 626-2 at 1 (Plaintiffs’ proposed verdict form, asking whether “ROSS infringed Thomson Reuters’ copyrights in Westlaw”). The scope of this alleged copyright is unspecified.

C. The Westlaw Platform and Its History

Westlaw.com first came online in 1998. Canter Decl. Ex. 9 (TRCC-00608822), at -829; *id.* Ex. 17 (TRCC-01314989), at -990. Its target audience is practicing lawyers at firms and within businesses, judges, and judicial clerks. *Id.* Ex. 18 (TR-0037447), at -456, -461. Before Westlaw came online, West made its collection of caselaw and copyrights available through books. *Id.* Ex. 9 (TRCC-00608822), at -825-27; *id.* Ex. 17 (TRCC-01314989), at -989 to 900.

Since its inception in the 1880s, the purpose of key numbers, headnotes, and the other materials that Plaintiffs call “Westlaw Content” has been to assist lawyers in finding judicial

opinions for legal research. Leiter Decl. ¶¶ 5-6, 10, 13-22, 25; Cox Decl. ¶ 8. West determined that the best way to achieve this was to organize cases in what is now called the key number system using the language of the law. West organized cases in this way so that lawyers could “look for precedents relating to specific subjects under heads to which they have become accustomed,” as he knew “any philosophical analysis of the law would be inapplicable to a digest[.]” Leiter Decl. ¶ 18.

Although Plaintiffs’ search platform is widely used, it is not available to everyone. Plaintiffs have admitted in interrogatory responses and requests for admission that they have a [REDACTED] See Canter Decl. Ex. 19 (Resp. to Req. for Admis. 72), at 46; *id.* Ex. 20 (Resp. to Interrog. 21), at 18 (Plaintiffs [REDACTED] [REDACTED]). And Plaintiffs have never taken any steps to try to sell their Westlaw Content as training data for artificial intelligence, to one of these legal research platform competitors, or anyone else. *Id.* Ex. 19 (Resp. to Req. for Admis. 72), at 45-46; *id.* Ex. 20 (Resp. to Interrog. 21), at 18; Cox. Decl. ¶¶ 11-12, 19-21.

1. Judicial Opinions are Available on Westlaw

Consumers can access judicial opinions on Westlaw. *See e.g.*, Canter Decl. Ex. 21 (TRCC-01928915). Plaintiffs describe its collection of judicial opinions as [REDACTED] and the [REDACTED] *Id.* Ex. 22 (TR-0432528), at -531. [REDACTED] *Id.* Ex. 23 (Resp. to Interrog. 26), at 25. Plaintiffs also [REDACTED] *Id.* Each page of a judicial opinion when printed states “Westlaw © [date] Thomson Reuters. No claim to original U.S. Government Works.” Canter Decl. Ex. 24 (TR-0812245), at -245.

2. Headnotes are Available on Westlaw

Headnotes have been available to West consumers for well over a century. *Id.* Ex. 20 (Resp. to Interrog. 15), at 9 [REDACTED] Canter Decl. Ex. 25 (TRCC-01390287), at -287.

Headnotes [REDACTED] [REDACTED] Canter Decl. Ex. 26 (L. Oliver Dep. Tr.), 32:21-25. Each point of law gets a headnote. *Id.* Ex. 27 (TR-0002864), at -872. The number one rule of the [REDACTED] for creating headnotes is that [REDACTED] [REDACTED] *Id.* Ex. 28 (TR-0040273), at -280; *id.* Ex. 27 (TR-0002864), at -872 [REDACTED] *id.* Ex. 8 (TRCC-00686613), at -627 (headnotes [REDACTED])

As of [REDACTED] headnotes in Westlaw, but [REDACTED] [REDACTED] so the number is now likely higher. *Id.* Ex. 20 (Resp. to Interrog. 14), at 7-8; *id.* (Resp. to Interrog. 16), at 10-11.

3. The Key Number System is Available on Westlaw

The West Key Number System was made publicly available in 1908. *See* Canter Decl. Ex. 9 (TRCC-00608822), at -826; *id.* Ex. 17 (TRCC-01314989), at -990. It was based on the American Digest System, a method for indexing the law first made available in 1887. *Id.* Ex. 9 (TRCC-00608822), at -826; Leiter Decl. ¶¶ 5-6. Over 75% of the topics in the key number system today are the same as the topics in the American Digest. *Id.* ¶ 23. And the American Digest was not original—it was based on preexisting casebooks and digests. *Id.* ¶¶ 5, 11-13.

The key number system is organized as follows: there are around 400 high-level topics, and each is given a number. Canter Decl. Ex. 29 (TR-0179830), at -834-36. Categorized under the topics are approximately 100,000 points of law, which are called “key numbers” or

subtopics, and each of these is also given a number. *Id.* Ex. 30 (TR-0523445), at -451; *id.* Ex. 31 (TR-0036398), at -407; *id.* Ex. 22 (TR-0432528), at -532; *id.* Ex. 33 (TR-0179843), at -843.

The topics and subtopics are dictated by the language of the law. They include “Criminal Law,” “Contracts,” and “Administrative Law and Procedure.” *Id.* Ex. 29 (TR-0179830), at -834. Because their purpose for lawyers is to find cases, the terms used in the key number system are terms lawyers are familiar with. *Canter Decl.* Ex. 26 (L. Oliver Dep. Tr.), 203:22-205:21. Topics and subtopics change as the law changes. For example, [REDACTED]

[REDACTED] *Id.* Ex. 26 (L. Oliver Dep. Tr.), 206:23-25. Topics have been added or removed as the law has evolved. *Leiter Decl.* ¶ 26; *Canter Decl.* Ex. 26 (L. Oliver Dep. Tr.), 202:21-203:2; *id.* 204:19-25 (noting that key numbers cover [REDACTED] and concepts of law); *see also id.* 202:7-203:10, 203:22-205:5. [REDACTED]

[REDACTED] *id.* Ex. 22 (TR-0432528), at -532.

4. Other “Editorial Enhancements” are Available on Westlaw

Other materials that Plaintiffs create that are on Westlaw include case synopses, notes of decision, and KeyCite. Case synopses [REDACTED]

[REDACTED] *Canter Decl.* Ex. 32 (TRCC-00631392), at -395. Notes of decision [REDACTED] *Id.* Plaintiffs describe KeyCite as [REDACTED]

[REDACTED] *Id.*

D. Plaintiffs’ Copyright in Westlaw Content

Plaintiffs’ copyrighted materials are registered as a compilation. *Canter Decl.* Ex. 34 (TR-0034557), at -558 This copyright covers the “original and revised text and compilation of legal material” prepared by the author, West. *Id.* The single copyright covers all the

copyrightable material on the Westlaw platform. *Id.* Ex. 35 (TR-0026188); *id.* Ex. 34 (TR-0034557). As of June 5, 2019, this compilation included 11,205,374,706 data records. *Id.* Ex. 35 (TR-0026188), at -190. As more material is “daily” added to the compilation, it is surely larger today. *Id.* at -189.⁴

E. The Search Technology Employed by Plaintiffs on Westlaw

Plaintiffs introduced a legal search tool on Westlaw in 2010 which it calls WestSearch. *Canter Decl.* Ex. 36 (TR-0433738); *id.* Ex. 37 (TR-0456869); *id.* Ex. 38 (TRCC-00618557), at -558; *id.* Ex. 9 (TRCC-00608822), at -830. Plaintiffs introduced a new legal search tool on Westlaw in 2018, four years after ROSS started—this tool is called WestSearch Plus. *Id.* Ex. 38 (TRCC-00618557), at -559.⁵

1. WestSearch – Algorithms Trained on User Behavior, Citation Networks, and Key Numbers; Searches All Legal Content; Results are Documents

A Westlaw user can search with the WestSearch tool by typing a query into the platform’s main search bar. *Canter Decl.* Ex. 36 (TR-0433738), at -738; *id.* Ex. 37 (TR-0456869), at -869.

After clicking search, [REDACTED] on the platform.

Canter Decl. Ex. 36 (TR-0433738), at -739. The [REDACTED]

[REDACTED],

⁴ Plaintiffs file a new registration with the U.S. Copyright Office every three months to reflect “updates” to this automated database. *See e.g.*, *Canter Decl.* Ex. 34 (TR-0034557), at -558. But it is still the same single copyright. *Id.*

⁵ Plaintiffs introduced WestSearch in 2010 on a version of their platform called Westlaw Next. *See Canter Decl.* Ex. 38 (TRCC-00618557) at -558 [REDACTED]

[REDACTED] Plaintiffs introduced WestSearch Plus on a version of their platform called Westlaw Edge. *See id.* at -559 ([REDACTED] Westlaw Edge includes both WestSearch and WestSearch Plus. *See id.* Ex. 1 (I. Moulinier Dep. Tr.), 30:15-21.

[REDACTED]

[REDACTED] *Id.*

A search across this [REDACTED] produces a list of the [REDACTED]
[REDACTED] *Id.* at -739, -741. The user can then filter down to obtain more specific results. *See id.*

At a programmatic level, WestSearch is comprised of two algorithms: the [REDACTED]
[REDACTED] and [REDACTED] *Id.*
Ex. 1 (I. Moulinier Dep. Tr.), 62:7-63:11; *id.* Ex. 39 (Expert Report of Jonathan Krein (Aug. 1, 2022) ¶ 88). Both algorithms are [REDACTED], which means that they are
[REDACTED]. *Id.* Ex. 39 (Expert Report of Jonathan Krein (Aug. 1, 2022) ¶ 88; Marks Decl. ¶ 15. Plaintiffs [REDACTED]
[REDACTED] Canter Decl. Ex. 1 (I. Moulinier Dep. Tr.), 92:1-9.

Plaintiffs have [REDACTED]. Canter Decl. Ex. 1 (I. Moulinier Dep. Tr.), 71:22-72:25; 88:14-17. [REDACTED]
[REDACTED] *id.* Ex. 9 (TRCC-00608822), at -830, and [REDACTED]
[REDACTED] *id.* Ex. 1 (I. Moulinier Dep. Tr.), 70:13-71:7, 88:14-17. The [REDACTED] and [REDACTED]
[REDACTED] included [REDACTED]
[REDACTED]; information on [REDACTED]
[REDACTED]; and [REDACTED] *Id.* Ex. 36 (TR-0433738), at -741.

WestSearch was also trained on [REDACTED]. Canter Decl., Ex. 1 (I. Moulinier Dep. Tr.), 126:11-127:8. Dr. Isabelle Moulinier, the Vice President of Applied Research for Artificial Intelligence at Thomson Reuters, was [REDACTED]. *Id.* 10:4-7; *id.*

11:5-18. She has stated that [REDACTED]

[REDACTED] in training WestSearch. *Id.* at 127:22-128:6.⁶

Finally, the [REDACTED] was [REDACTED]

Canter Decl. Ex. 1 (I. Moulinier Dep. Tr.), 73:15-74:14. [REDACTED]

[REDACTED]. *Id.*

74:15-75:1. The questions in these question-document pairs came from [REDACTED]

[REDACTED] *Id.* 75:24-76:6.

Headnotes on their own were not used as the questions or the answers for the WestSearch training data. *Id.* 76:21-77:16.

2. WestSearch Plus – Trained with Headnotes as the *Answers* in QA Pairs; Only Responds to Questions that Plaintiffs Pose; Results are Headnotes

Like WestSearch, a user begins her search with the WestSearch Plus tool by typing text into the main Westlaw search bar. Canter Decl. Ex. 40 (TR-0477495), at -500; *id.* Ex. 41 (TRCC-00218156), at -157. However, a WestSearch Plus search only occurs if the user selects one of the ‘suggestions’ that pop-up below the search bar after the query is drafted. *See id.* Ex. 42 (TRCC-01012442), at -442; *id.* Ex. 43 (TRCC-01013084), at -089; *id.* Ex. 67 (TRCC-00992673), at -690 (WestSearch Plus [REDACTED] *id.* Ex. 1 (I. Moulinier Dep. Tr.), 15:6-11 (the Westlaw feature [REDACTED]

[REDACTED]⁷

⁶ Features are described below. *See* Statement of Facts § L.3.

⁷ [REDACTED]

[REDACTED] Canter Decl. Ex. 1 (I. Moulinier Dep. Tr.), 30:15-21.

WestSearch Plus searches are only based on questions that WestSearch Plus poses.

[REDACTED] no WestSearch Plus suggestions are generated based on the user's query, which means that the customer frequently cannot use WestSearch Plus to answer her legal question.

Canter Decl. Ex. 44 (TRCC-00988301), at -313.

When a user can actually complete a search with WestSearch Plus, the answers come from Plaintiffs' corpus of headnotes. Canter Decl. Ex. 45 (TR-0908443), at -443; *id.* Ex. 46 (Al-Kofahi Dep. Tr.), 41:12–15; *id.* Ex. 39 (Expert Report of Jonathan Krein (Aug. 1, 2022)), ¶ 86.

At a programmatic level, WestSearch Plus [REDACTED]

[REDACTED] Canter Decl. Ex. 47 (TRCC-00194937), at -953. [REDACTED]

[REDACTED] which is the [REDACTED]

[REDACTED] *Id.* Ex. 39 (Expert Report of Jonathan Krein (Aug. 1, 2022)) ¶ 60.

The [REDACTED]

[REDACTED] Canter Decl. Ex. 45 (TR-0908443), at -443. The [REDACTED]

[REDACTED]. *Id.* The [REDACTED]. *Id.*; *id.* Ex. 39 (Expert

Report of Jonathan Krein (Aug. 1, 2022)) ¶ 79; *id.* Ex. 1 (I. Moulinier Dep. Tr.), 77:18-78:4.

F. To Develop Its Legal Research Tool, ROSS Needed Pairs of Legal Questions and Judicial Quotes that Answer those Legal Questions; It Got This in the Form of Memoranda

To develop its legal search tool, ROSS needed a large and diverse body of legal questions and answers to those questions that were verbatim quotes from judicial opinions. Ovbiagele Decl. ¶ 30; Canter Decl. Ex. 48 (TR-0000568), at -569.

ROSS obtained its collection of legal question and answer pairs in the form of memoranda. *See e.g.*, Ovbiagele Decl. Ex. 49 (ROSS-000000001) (an example memo). The memos contained one legal question and four to six quotations from a judicial opinion that were answers to that question. *See e.g.*, Canter Decl. Ex. 50 (C. Cahn Dep.) 54:15-55:21.

The questions were legal questions that a lawyer would ask. *See e.g.*, Ovbiagele Decl. Ex. 49 (ROSS-000000001), at -001; Canter Decl. Ex. 48 (TR-0000568), at -568 (“Each Memo shall include a Legal Research Question”). In addition, the question-answer pairs, needed to be “representative of the range of language and expression that users of the ROSS system could be expected to pose to the system.” Ovbiagele Decl. ¶ 30.

The answers were quotes precisely as they appear in legal opinions. *See* Ovbiagele Decl. Ex. 49 (ROSS-000000001), at -001-02 (listing verbatim paragraphs from judicial opinions under “Case” 1 to 6); Canter Decl. Ex. 48 (TR-0000568), at -568. ROSS needed the answers to be quotations from judicial opinions because the output of ROSS’ legal search tool is quotations from judicial opinions. Ovbiagele Decl. Ex. 11 (ROSS-023179081), at -087 (“ROSS returns the exact passages from the law that answer their question.”).⁸

Each quote was given a relevancy rating—great, good, topical or irrelevant. *See e.g.*, Ovbiagele Decl. Ex. 49 (ROSS-000000001), at -001-02; Canter Decl. Ex. 48 (TR-0000568), at -568-69 (sections 5.3 through 5.6 describing the relevancy labels). A “great” quote was “one that contain[ed] an answer to all essential elements” of the legal research question. *Id.* Ex. 48 (TR-0000568), at -568 (section 5.3).⁹

G. LegalEase Prepared the Memos Containing the Question-Answer Pairs

ROSS hired a third-party legal research company, LegalEase Solutions, LLC, to prepare the memos. Canter Decl. Ex. 2 (J. Ovbiagele Dep. Tr.), 116:12-16; *id.* Ex. 48 (TR-0000568). In turn, LegalEase hired a subcontractor Morae Global (collectively “LegalEase”) to prepare the

⁸ This is also why [REDACTED]. Canter Decl. Ex. 45 (TR-0908443) at -443; *id.* Ex. 46 (Al-Kofahi Dep. Tr.), 41:12–15.

⁹ The AI legal search engine needed to learn great quotes and also quotes that ran the gamut to irrelevant. As is set forth below, the AI legal search engine needed to know the mathematical relationship of words, including words that would not match a query. Statement of Facts § L.3-5.

memos. *Id.* Ex. 51 (T. Hafeez Dep. Tr.), 88:13-18; *id.* Ex. 50 (C. Cahn Dep. Tr.), 25:2-28:1; *id.* Ex. 52 (TR-0039943) (the agreement between LegalEase and Morae).

LegalEase created approximately 25,000 memos between September 2017 to December 2017. *Id.* Ex. 53 (TR-0039043) at 124:10-11. The project through which these memos were created was called the “Bulk Memo Project.” Canter Decl. Ex. 54 (T. van der Heijden Dep. Tr.), 321:7-322:3; *id.* Ex. 48 (TR-0000568).

H. How LegalEase Prepared the Memos and What LegalEase Gave ROSS

LegalEase used Westlaw and LexisNexis (a competitor legal research platform) to prepare the memos. Canter Decl. Ex. 55 (T. Whitehead Dep. Tr.), 60:10-19. ROSS did not direct LegalEase to use Westlaw. *Id.* 43:2-5.¹⁰

By accessing Westlaw, LegalEase made ephemeral copies of any copyrighted content to access judicial opinions. This occurred because there is no way to access the judicial opinions on the platform unless a user also accesses the copyrighted content as well. Canter Decl. Ex. 29 (TR-0179830), at -830 (“*Each legal issue in a case published by Thomson Reuters is identified, summarized in a headnote, and assigned a topic and key number in the West Key Number System.*”) (emphasis added).

The memos were prepared in word documents. Canter Decl. Ex. 55 (T. van der Heijden Dep. Tr.), 133:7-19. The filenames for these word documents did not include key numbers or any tangible representation of the key number system. *Id.* Ex. 56 (Frederiksen-Cross Dep. Tr.), 343:6-21; Barbara Frederiksen-Cross Decl., ¶¶ 27. LegalEase added legal topic information to

¹⁰ LegalEase paid to use Westlaw pursuant to a license. Canter Decl. Ex. 51 (T. Hafeez Dep. Tr.), 321:3-19. That agreement has a Fair Use provision. *Id.* Ex. 57 (TR-002808) at -808 (section 1.a.(iii)); D.I. 16, 16-1 at Ex. 1.

the file names for its own internal tracking purposes. Canter Decl. Ex. 2 (J. Ovbiagele Dep. Tr.), 132:7-15; *id.* Ex. 55 (T. Whitehead Dep. Tr.), 101:2-10, 107:6-108:9.

I. To Complete the Memos, LegalEase Drafted the Legal Questions to Be as Close as Possible to the Language in Judicial Opinions and Took Verbatim Paragraphs from Judicial Opinions for the Answers

When preparing the memos, LegalEase contract attorneys logged into Westlaw (or Lexis), located a topic in the key number system, and clicked through the topic and a subtopic to arrive at a headnote associated with a judicial opinion—just like anyone else could do who completes legal research with the key number system. Canter Decl. Ex. 58 (LEGALEASE-00078065), at -067-76 (the [REDACTED] describing the steps to prepare memos using Westlaw and Lexis); *id.* Ex. 51 (T. Hafeez Dep. Tr.), 50:9-23.

Then, the contract attorney drafted a question based on the judicial opinion. *Id.* Ex. 58 (LEGALEASE-00078065), at -067-76. While the question drafting started with the headnote, the contract attorneys [REDACTED]

[REDACTED] *Id.* Ex. 50 (C. Cahn Dep. Tr.), 64:22-65:13.

In fact, the contract attorneys were directed to not copy the headnotes. The [REDACTED] that contract attorneys followed for memo drafting warned that [REDACTED] *Id.* Ex. 58 (LEGALEASE-00078065), at -068. The contract attorneys were instructed not to copy headnotes into questions. *Id.* Ex. 51 (T. Hafeez Dep. Tr.), 63:6-25, 78:1-17, 84:20-85:9, 225:20-228:24; *id.* Ex. 59 (R-LEGALEASE-00050673), at -673 (LegalEase directing contract attorneys to [REDACTED] [REDACTED]); *id.* Ex. 55 (T.

Whitehead Dep. Tr.), 127:5-18; *id.* Ex. 50 (C. Cahn Dep. Tr.), 62:21-63:21, 64:22-65:13. And LegalEase did not want its workers to copy headnotes. *Id.* Ex. 60 (TR-0047926), at -926 (LegalEase stating that [REDACTED])

because [REDACTED]

[REDACTED]).

With regard to the quotes, each quote needed to accurately represent one of the four relevance rankings. *Id.* Ex. 48 (TR-0000568), at -568-69; *id.* Ex. 50 (C. Cahn. Dep. Tr.), 132:23-133:7. The contract attorneys searched for a quote that accurately reflected the rating, copy and paste the quote into the memo, and then add the appropriate rating as well. *Id.* Ex. 58 (LEGALEASE-00078065), at -077-78.

The use of headnotes as described here was no substitute for locating the requisite quotes from the judicial opinions. *Id.* Ex. 50 (C. Cahn Dep. Tr.), 69:8-71:14.

J. The Completed Memos that LegalEase Gave to ROSS Did Not Contain Plaintiffs' Copyrighted Content

The memos did not include key numbers, case synopses, any tangible representation of the key number system, or any other possible “editorial enhancements” as Plaintiffs define the term. Ovbiagele Decl. Ex. 49 (ROSS-000000001); Canter Decl. Ex. 50 (C. Cahn Dep. Tr.), 125:7-12 [REDACTED] Marks Decl. ¶ 34; Ovbiagele Decl. ¶¶ 21, 33. This is apparent on the face of the memos. *See, e.g.*, Ovbiagele Decl. Ex. 49 (ROSS-000000001). ROSS never sold or licensed these memos and never attempted to do so either. Ovbiagele Decl. ¶¶ 41-44; Cox Decl. ¶¶ 13, 25.

K. The Additional Items Provided to ROSS by LegalEase that ROSS Rejected

1. List of 91 Legal Topics

As part of ROSS’s ongoing efforts to improve its search tool, ROSS ran experiments to determine whether it should use “classifier” technology. Ovbiagele Decl. ¶ 36-37; Canter Decl. Ex. 54 (T. van der Heijden Dep. Tr.), 415:14-21. ROSS intended that the classifier technology would classify judicial opinions into legal categories, such as civil procedure or bankruptcy,

before being searched. Ovbiagele Decl. ¶¶ 36-37; Canter Decl. Ex. 54 (T. van der Heijden Dep. Tr.), 415:14-21. However, ROSS's classification experiments did not produce results that led to ROSS deploying new classification technology. Ovbiagele Decl., ¶ 38-40; Canter Decl. Ex. 54 (T. van der Heijden Dep. Tr.), 415:22-416:10.

[REDACTED]

[REDACTED]

[REDACTED] Ovbiagele Decl. ¶ 39-40; Ovbiagele Decl. Ex. 61 (ROSS-010128683). ROSS's list of 38 topics was premised on [REDACTED]

[REDACTED] not on the list from LegalEase. Canter Decl. Ex. 62 (T. van der Heijden Jan. 28, 2020 Dep. Tr.), 76:14-19 [REDACTED]

[REDACTED]

Because ROSS did not pursue any classification projects for its search tool, it did not make it into ROSS's search engine. *Id.* Ex. 1 (J. Ovbiagele Dep. Tr.), 131:8-17, 133:5-6; Ovbiagele Decl. ¶ 40.

2. 500 Judicial Opinions

In December 2017, LegalEase sent 500 judicial opinions to ROSS via email. Ovbiagele Decl. Ex. 63 (ROSS-000197949), at -949-50. ROSS responded by telling LegalEase that it did not want any of the headnotes, key numbers, or other copyrighted content when receiving those cases—it just wanted [REDACTED] *Id.* at -949. ROSS said that it was [REDACTED] and [REDACTED]

[REDACTED]

[REDACTED] Canter Decl. Ex. 54 (T. van der Heijden Dep. Tr.), 135:23-136:23.

L. The Development of ROSS's Legal Research Tool and How it Operates

To develop its legal search engine, ROSS used the following process.

1. ROSS First Extracted the Question-Answer Pairs from the Memos, Discarded Some of the Pairs, and Randomly Separated the Pairs for Different Uses

First, ROSS [REDACTED]

[REDACTED] Ovbiagele Decl. ¶¶ 16-17; Marks Decl. ¶ 12. During this process of [REDACTED]

[REDACTED] ROSS [REDACTED]

[REDACTED] whether the relevance ratings were accurate. Ovbiagele Decl.

¶¶ 7-9, 16-17; Marks Decl. ¶ 12. Question-answer pairs with [REDACTED]

[REDACTED] *Id.* Each [REDACTED]

[REDACTED]. Ovbiagele Decl. ¶ 17.

Additional question-answer pairs were [REDACTED]. Ovbiagele

Decl. ¶¶ 9-10, 16-17; Marks Decl. ¶ 12. If the case from which [REDACTED]

[REDACTED]. *Id.* ROSS's

repository of cases was licensed from a company called Casemaker. Ovbiagele Decl. ¶ 14; *id.*

Ex. 64 (ROSS-009721314). The judicial opinions from Casemaker did not contain any Westlaw Content. *Id.*

ROSS randomly allocated [REDACTED] of the remaining question-answer pairs for use in training and [REDACTED] for use in testing and validation. Ovbiagele Decl. ¶¶ 6, 32; Marks Decl. ¶ 12.¹¹ Because the pairs used for training, testing, and validation were randomly selected, any order or organization across the collection of pairs that previously existed was eliminated. Marks Decl. Ex. 12; Canter Decl. Ex. 39 (Expert Report of Jonathan Krein (Aug. 1, 2022) ¶ 52, n.23

¹¹ [REDACTED]

[REDACTED] Canter Decl. Ex. 39 (Expert Report of Jonathan Krein (Aug. 1, 2022)) ¶ 50.

(acknowledging that data is [REDACTED] for testing, training, and validation);
Ovbiagele Decl. ¶ 32.

2. ROSS Then Prepared the Question-Answer Pairs for Computational Analysis of the Text's Non-Expressive Characteristics

ROSS then prepared the question-answer pairs to become training data. Ovbiagele Decl. ¶ 18; Marks Decl. ¶¶ 17-19. [REDACTED]

[REDACTED] *Id.* [REDACTED]
[REDACTED] Ovbiagele Decl. ¶ 18; Marks Decl. ¶¶ 18-19. Tokenizing text means that ‘stop’ words were removed. *Id.* These words—for example, “is,” “to,” “be,” “of,” “by”—are irrelevant for computational analysis. *Id.* For example, the phrase “is involuntary manslaughter considered to be a lesser offense of murder” becomes, after tokenization, “involuntary manslaughter considered lesser offense murder.” Marks Decl. ¶ 19.

[REDACTED]. Ovbiagele Decl. ¶ 18; Marks Decl. ¶¶ 18-19. This means changing all of the grammatical inflections to the same form—“swam,” “swum,” and “swimming” all become “swim.” *Id.* Using the same example, after lemmatization, the above phrase becomes “involuntary manslaughter consider lesser offense murder.” Marks Decl. ¶ 19.

3. ROSS Featurizes the Question-Answer Pairs—Extracting Only Numerical Representations of the Relationships Among the Language and Discarding the Actual Text

The tokenized and lemmatized data was then reduced to mathematical equations. The machine-learning model was not trained on the text of the questions and answers. Ovbiagele Decl. ¶¶ 4, 19, 22-30; Marks Decl. ¶¶ 13-16, 20-24, 29-36. [REDACTED]

[REDACTED]. Ovbiagele Decl. ¶¶ 19, 22-30;

Marks Decl. ¶¶ 20-21, 23, 29-36. Rather than extracting the natural language expression through this process, measurements about the text and statistic relationship between the text are extracted. *Id.* This process is called *featurization*. Ovbiagele Decl. ¶ 19; Branting ¶¶ 20. [REDACTED]
[REDACTED] See Ovbiagele Decl. ¶ 19; Marks Decl. ¶ 20.

ROSS programmatically transformed each one of its remaining question-answer pairs into [REDACTED] separate features. See Ovbiagele Decl. ¶ 19; Marks Decl. ¶ 20.¹² Most of these features generated a number that reflected something about the relationship between the question and answer— [REDACTED]

[REDACTED] Ovbiagele Decl. ¶ 19.20; Marks Decl. ¶ 20.20. [REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

See Ovbiagele Decl. ¶ 19.24 to .27; Marks Decl. ¶ 20.24 to .27. [REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

[REDACTED] See Ovbiagele Decl. ¶ 19.14; Marks Decl. ¶ 20.14.

In every case, the feature extracted information out of the question-answer pair that was not the same as the natural language meaning of the text in that question-answer pair. Ovbiagele

¹² [REDACTED]

[REDACTED] See Marks Decl. ¶¶ 25-28.

Decl. ¶¶ 28-29; Marks Decl. ¶¶ 21, 23, 29. This extraction was so extreme that the text of the pairs cannot be rediscovered after the numbers are generated. Marks Decl. ¶ 23.

[REDACTED]
[REDACTED]. Ovbiagele Decl. ¶ 20; Marks Decl. ¶ 21. [REDACTED]

[REDACTED]. Ovbiagele Decl. ¶¶ 23-28; Marks Decl. ¶¶ 14-16, 21.

4. ROSS Uses the Numbers that Represent Abstract Characteristics of the Text to “Train” the Machine-Learning Model

As part of the training process—the process of turning mathematical relations about text into a function that runs when code commands are issued—ROSS used a machine-learning model called [REDACTED]. Ovbiagele Decl. ¶¶ 22-28; Marks Decl. ¶ 15-16, 21. The numerical features are used by the “learning to rank” component of [REDACTED] to compute parameters (or guideposts) that train the technology to predict whether a candidate passage is relevant. Ovbiagele Decl. ¶ 26; Marks Decl. ¶ 15. The relevance labels associated with the features, “great,” “good,” “topical,” and “irrelevant,” contribute to this training because they point the learning technology in the right direction. Ovbiagele Decl. ¶ 25; Marks Decl. ¶¶ 10, 15. Through this process, the features, by setting the parameters for the ranking algorithm, are causing that component of the technology to change. Ovbiagele Decl. ¶ 26; Marks Decl. ¶ 15-16.

The actual answer to great quotes in the memos were not preserved or used to train the model at all, and thus ROSS’s tool would not and could not simply return what was the great quote found in the memo based on the memo’s legal question being asked. Ovbiagele Decl. ¶¶ 22, 32-35; Marks Decl. ¶¶ 15-16, 20-22. Moreover, any order, structure, or organization of the numerical features was, at this training stage, even further abandoned, because the model did not

preserve the order of the numbers on which it was trained. Ovbiagele Decl. ¶¶ 32-35; Marks Decl. ¶¶ 30-33.

Once the features had been derived, ROSS did not use the text of the question-answer pairs or the memos, Ovbiagele Decl. ¶¶ 22; Marks Decl. ¶ 22, and the question-answer pairs were discarded. Ovbiagele Decl. ¶¶ 22; Marks Decl. ¶ 22.¹³ Dr. Moulinier stated the same point about WestSearch Plus question-answer pairs—that after the search tool is [REDACTED] it is [REDACTED] and the question-answer pairs are no longer used or relevant. Canter Decl. Ex. 1 (I. Moulinier Dep. Tr.), 61:22-62:6.

5. How ROSS's Search Tool Operates When a User Makes a Query – Users Can Ask Any Questions; the Outputs are Only Judicial Quotes

A user of ROSS's search tool benefits from ROSS's innovations in the legal research market. The innovations are clearest when compared to Plaintiffs' tools.

Unlike WestSearch, ROSS's tool only searches for paragraphs from judicial text, and relies on no headnotes or other human-mediated material. Unlike WestSearch, ROSS's tool is only trained on the characteristics of the judicial passages and legal questions and is not trained on material that is only indirectly related to the judicial text, such as [REDACTED]

[REDACTED] *Compare* Statement of Facts § E.1 *with* L.

Unlike WestSearch Plus, ROSS's tool responds to every natural language question a user asks. Unlike WestSearch Plus, the results from a search using ROSS's tool is the actual text of a judicial opinion that directly answers the question, not a human-created headnote that itself may not reflect the judicial opinion. *Compare* Statement of Facts § E.2 *with* L.

¹³ There is no evidence that LegalEase maintained the memos—in fact they were directed to destroy them—or the results of any of the searches required to create the memos Canter Decl. Ex. 48 (TR-0000568) at -570-71 (section 13).

ROSS has thus provided users with a new way to complete legal research that was not previously available. The user inputs any natural language question into ROSS’ search bar. *See* Ovbiagele Decl. Ex. 13 (ROSS-009542437), at -437. The tool then searches across ROSS’s repository of millions of judicial opinions. *Id.* ¶¶ 13-14; Marks Decl. ¶¶ 8-9, 24, 35.

The output was a list of paragraphs from judicial opinions with the language of Judges directly answering the question that the user has asked. *See* Ovbiagele Decl. Ex. 11 (ROSS-023179081), at -087; Marks Decl. ¶ 35. It included no Westlaw Content. Ovbiagele Decl. Ex. 12 (ROSS-003515560) at -560; *id.* ¶ 14.

ARGUMENT

“The ultimate goal of copyright is to expand public knowledge and understanding . . . Thus, while authors are undoubtedly important intended beneficiaries of copyright, the ultimate, primary intended beneficiary is the public, whose access to knowledge copyright seeks to advance by providing rewards for authorship.” *Authors Guild v. Google, Inc.*, 804 F.3d 202, 212 (2d Cir. 2015). Because giving authors completely unfettered and unlimited control over any copying of their works would undermine this goal, “[f]rom the infancy of copyright protection, some opportunity for fair use of copyrighted materials has been thought necessary to fulfill copyright’s very purpose.” *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 575 (1994).

In determining fair use, courts examine four factors: “(1) the purpose and character of the use . . . (2) the nature of the copyrighted work; (3) the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and (4) the effect of the use upon the potential market for or value of the copyrighted work.” 17 U.S.C. § 107.

First, we discuss the factors that assess the copyrights—factors 2 and 3. Then, we move onto the factors that assess the infringing use of the copyrights—factors 1 and 4. Each of the four factors favors a finding of fair use.

I. FACTOR 2: NATURE OF THE COPYRIGHTED WORK

The second fair use factor examines the nature of the copyrighted work. A work can only be creative, derivative, or compiled. *See* 17 U.S.C. § 102 (creative); *id.* § 103 (derivative and collective). Creative works receive the most protection, then derivative works, and finally compiled works. *Intervest Constr., Inc. v. Canterbury Est. Homes, Inc.*, 554 F.3d 914, 919 n.3 (11th Cir. 2008). Additionally, “factual,” “utilitarian,” or “informational,” works receive less protection than creative works because they are farther from copyright’s core. *Campbell*, 510 U.S. at 586; *Hustler Mag. Inc. v. Moral Majority Inc.*, 796 F.2d 1148, 1153–54 (9th Cir. 1986); 4 Melville B. Nimmer & David Nimmer, NIMMER ON COPYRIGHT § 13F.06[A] (rev. ed. 2024) (“Nimmer on Copyright”) (“[T]he more informational or functional is plaintiff’s work, the broader should be the scope of the fair use defense.”).

Plaintiffs’ copyrights are factual, functional, and informational, not creative like a novel or original song. They are intended to reflect and essentially report on the text of judicial opinions so that lawyers can use and find them. *See* Statement of Facts § C. They are purposefully constrained by the language of judicial opinions and of the law. *Id.* That is a hard limit, *i.e.*, a constraint, on how creative Plaintiffs’ work could possibly be. *Matthew Bender & Co. v. W. Publ’g Co.*, 158 F.3d 674, 682 (2d Cir. 1998) (“The creative spark is missing where: (i) industry conventions or other external factors so dictate selection that any person composing a compilation of the type at issue would necessarily select the same categories of information ... (ii) the author made obvious, garden-variety, or routine selections...” (internal citations omitted); *Mitel, Inc. v. Iqtel, Inc.*, 124 F.3d 1366, 1375 (10th Cir. 1997) (“We have extended this traditional copyright doctrine to exclude from protection against infringement those elements of a work that necessarily result from external factors inherent in the subject matter of the work.”);

Dun & Bradstreet Software Servs., Inc. v. Grace Consulting, Inc., 307 F.3d 197, 215 (3d Cir. 2002) (discussing *scènes à faire*); *see also* Statement of Facts § C.3; Leiter Decl. ¶ 26.

As Plaintiffs readily admit, the key number system is used to help navigate its database and locate cases. *See* Statement of Facts § C.3. In other words, the value of the key number system is functional, thus making fair use more likely. *See Google LLC v. Oracle Am., Inc.*, 593 U.S. 1, 29 (2021).¹⁴

With headnotes, the same holds true. *See* Statement of Facts § C.2. Whether treated as part of a compilation or as stand-alone works, headnotes are intended to allow lawyers to locate case law. *Id.* Unbounded creativity is eschewed by Plaintiffs. *Id.* Every point of law receives a headnote. *Id.* And “Generally, the editor should follow the court’s language when creating headnotes.” Canter Decl. Ex. 28 (TR-0040273), at -280. Plaintiffs even admit in their internal documents that headnotes [REDACTED] Canter Decl. Ex. 8 (TRCC-00686613), at -627.

When the purpose of the work is to deliver factual information, even if some original expression (or creativity) is added, the copyright protection extended is thin. *See e.g., Monsarrat v. Newman*, 28 F.4th 314, 323 (1st Cir. 2022) (quoting from a website’s policy in a blog post was minimally creative); *SOFA Ent., Inc. v. Dodger Prods., Inc.*, 709 F.3d 1273, 1279 (9th Cir. 2013) (use of short clip from show to convey factual information favored finding of fair use); *Swatch Grp. Mgmt. Servs. Ltd. v. Bloomberg L.P.*, 756 F.3d 73, 89 (2d Cir. 2014) (finding that factor two favored fair use when a news company copied the sound recording of a company’s earnings

¹⁴ While “[f]actor two analysis typically focuses on the divide between the creative and the factual rather than that between the creative and the functional or utilitarian . . . because defendants conventionally assert that a work’s functional features undermine the work’s copyrightability . . . Their frequent success on that basis has pretermitted consideration of the analytically subsequent issue of fair use.” Nimmer on Copyright § 13F.06.

report because “assum[ing] ... that the call contained sufficient original expression—in the form of the executives’ tone, cadence, accents, and particular choice of words—to be copyrightable, the purpose of the call was not ... to showcase those forms of expression. Rather, the call’s sole purpose was to convey financial information”); *see also Harper & Row Publishers, Inc. v. Nation Enterprises*, 471 U.S. 539, 563 (1985) (“The law generally recognizes a greater need to disseminate factual works than works of fiction or fantasy.”).

Finally, key numbers and headnotes all fall within a compilation. *Matthew Bender & Co.*, 158 F.3d at 707; *see also, e.g., Miller v. Universal City Studios, Inc.*, 650 F.2d 1365, 1368-69 and 1372 (5th Cir. 1981) (use of facts from one work made author only a “discoverer” of materials and the work is only protectable as a compilation); *S. Credentialing Support Servs., L.L.C. v. Hammond Surgical Hosp., L.L.C.*, 946 F.3d 780, 783 (5th Cir. 2020) (“the selection and arrangement of information that enable effective use” is copyrightable but that copyright is “necessarily thin”); *Apple Barrel Prods., Inc. v. Beard*, 730 F.2d 384, 387 (5th Cir. 1984) (collecting cases).¹⁵ As such, the protection extended is thin. *See, e.g., Craft Smith, LLC v. EC Design, LLC*, 969 F.3d 1092, 1103-04 (10th Cir. 2020); *Experian Info. Sols., Inc. v. Nationwide Mktg. Servs. Inc.*, 893 F.3d 1176, 1186 (9th Cir. 2018); *BellSouth Advert. & Pub. Corp. v. Donnelley Info. Pub., Inc.*, 999 F.2d 1436, 1445 (11th Cir. 1993).¹⁶

¹⁵ When a compilation includes original works, it is a collective work, the copyright in that collective work does not enlarge or contract the scope of copyright in the underlying works that comprise it. Nimmer on Copyright § 3.04. The underlying works must be analyzed on their own terms as either creative or derivative and not as the compilation.

¹⁶ ROSS understands that individual headnotes may receive copyright protection standing apart from the compilation. It is not clear that Plaintiffs continue to urge that headnotes were infringed outside of their place in a compilation. *See* Canter Decl. Ex. 65 (Aug. 6 Pre-Trial Tr.), 29:19-20 (“This isn’t a case about matching language.”). However, even if taken individually, as noted above, headnotes would still only receive thin copyright protection.

Because the Westlaw Content is factual and functional in nature, this factor weighs in ROSS's favor.

II. FACTOR 3: AMOUNT AND SUBSTANTIALITY OF THE WORK TAKEN

Factor three of the fair use analysis considers the amount and substantiality of the work taken. “Substantively, the third factor considers ‘whether the quantity and value of the materials used, are reasonable in relation to the purpose of the copying.’” *Whiddon v. Buzzfeed, Inc.*, 638 F. Supp. 3d 342, 354 (S.D.N.Y. 2022) (quoting *Blanch v. Koons*, 467 F.3d 244, 257 (2d Cir. 2006)); *see also Nihon Keizai Shimbun, Inc. v. Comline Bus. Data, Inc.*, 166 F.3d 65, 73 (2d Cir. 1999) (“In applying the third factor, ‘what is relevant is the amount and substantiality of the copyrighted expression that has been used, not the factual content of the material in the copyrighted works.’” (quoting *Salinger v. Random House, Inc.*, 811 F.2d 90, 97 (2d Cir. 1987))).

ROSS “took” an insignificant portion of the copyrights at issue.¹⁷ At most, ROSS infringed the text of [REDACTED] headnotes, and the relationship between [REDACTED] headnotes and case passages. *See* Statement of Facts § B.1-2. This is less than 0.1% of the total number of headnotes that Plaintiffs possess, and less than 0.1% of the total number of headnote-case passage relationships that Plaintiffs possess. *See id.* § C.2 ([REDACTED]).¹⁸

As for key numbers and key number system, there are [REDACTED] topics and subtopics in the key number system. *See id.* § C.3. Plaintiffs allege that LegalEase used 5,473 of the subtopics, or at most, 5.473% of the key number system. *See id.* § B.3. As for the [REDACTED] cases that

¹⁷ The numbers contained in this section cannot include what Plaintiffs claim LegalEase infringed. Plaintiffs did not disclose the number of key numbers, headnotes, or any other portion of Westlaw Content LegalEase infringed. They should not be allowed to do so now.

¹⁸ The math is [REDACTED] (total # of infringed HNs) / [REDACTED] (total # of HNs) = 0.00075966596.

LegalEase sent to ROSS (and which ROSS did not want), *see id.* § B.5, that equals less than 0.001% of the [REDACTED] cases in Westlaw. *See id.* § C.1.¹⁹

The numbers are even smaller when compared to the actual work. Plaintiffs state that its Westlaw compilation contains 11,205,374,706 data records. *See id.* § D. Adding up all the headnotes, topics, subtopics, and portions of the key number system equals .000444% of what Plaintiffs have registered. *See id.* §§ B-C.²⁰

This minute percentage of a factual, utilitarian, and informational compilation is fair use. *Matthew Bender*, 158 F.3d at 678 (“West conceded (immediately prior to trial) [what] would be permissible under the fair use doctrine, *i.e.*, one to two percent of its published reports of Supreme Court and court of appeals cases.”); *see also Google*, 141 S. Ct. at 1204-06 (copying 0.4% of work fair use); *New Era Publ’ns Int’l, ApS v. Carol Publ’g Grp.*, 904 F.2d 152, 158-59 (2d Cir. 1990) (factor favored fair use when defendant used a “small percentage” of the copyrighted work.”).

Not only did ROSS take an insignificant portion of Plaintiffs’ work, but its use would be fair even if it had taken far more. “Complete unchanged copying has repeatedly been found justified as fair use when the copying was reasonably appropriate to achieve the copier’s transformative purpose and was done in such a manner that it did not offer a competing substitute for the original.” *Authors Guild*, 804 F.3d at 221; *see also Whiddon v. Buzzfeed, Inc.*, 638 F. Supp. 3d at 354 (“A secondary work may copy the original in its totality where a full copy is reasonably appropriate to the secondary work’s transformative purpose.”).

¹⁹ The math here is [REDACTED] infringed cases / [REDACTED] total cases = 0.0000357.

²⁰ The math here is the following: [REDACTED] infringed cases + [REDACTED] infringed headnotes + [REDACTED] infringed headnote-case-passage relationships + [REDACTED] topics and subtopics equals 49,727. And 49,727 / 11,205,374,706 equals 0.00000443778.

III. FACTOR 1: THE PURPOSE AND CHARACTER OF THE USE

The first fair use factor is the purpose and character of the use. The question is whether the new work “supersedes the objects of the original creation, or instead adds something new, with a further purpose or different character, altering the first with new expression, meaning, or message; it asks, in other words, whether and to what extent the new work is ‘transformative.’” *Sony Comput. Entm’t, Inc. v. Connectix Corp.*, 203 F.3d 596, 602-08 (9th Cir. 2000) (quoting *Campbell v. Acuff-Rose Music, Inc.*, 510 U.S. 569, 579 (1994)). “[T]he more transformative the [secondary] work, the less will be the significance of other factors, like commercialism, that may weigh against a finding of fair use.” *Campbell*, 510 U.S. at 579. “The use of a copyrighted work need not alter or augment the work to be transformative in nature. Rather, it can be transformative in function or purpose without altering or actually adding to the original work.” *A.V. ex rel. Vanderhye v. iParadigms, LLC*, 562 F.3d 630, 639 (4th Cir. 2009); *see also Authors Guild, Inc. v. HathiTrust*, 755 F.3d 87, 96 (2d Cir. 2014) (“Added value or utility is not the test: a transformative work is one that serves a new and different function from the original work and is not a substitute for it.”).

Courts have found to be transformative far less than what ROSS did here. For example, in *HathiTrust*, the Second Circuit found that merely digitizing text to allow full-text searching was transformative. 755 F.3d at 97. In *Authors Guild*, the same court found that displaying “snippets” of text from books to allow researchers to see if that book was relevant was a transformative use. 804 F.3d at 217-18. And in *White v. W. Publ’g Corp.*, West Publishing convinced a court that the wholesale copying of briefs for the purposes of allowing lawyers to conduct research was transformative. 29 F. Supp. 3d 396, 399 (S.D.N.Y. 2014).

Other cases show just how far beyond the transformative standard ROSS’s use was. ROSS did more than create a video backdrop for live music performances using a copyrighted

photograph. *Seltzer v. Green Day, Inc.*, 725 F.3d 1170, 1177 (9th Cir. 2013). It did more than collect images of posters for inclusion in a historical work, *Bill Graham Archives v. Dorling Kindersley Ltd.*, 448 F.3d 605, 612-13 (2d Cir. 2006), or make a collage using photographs, *Cariou v. Prince*, 714 F.3d 694, 706 (2d Cir. 2013),²¹ or use a commercial image in a mixed-media painting, *Blanch v. Koons*, 467 F.3d 244, 252–53 (2d Cir. 2006), or display thumbnails of images from other websites. *Kelly v. Arriba Soft Corp.*, 336 F.3d 811, 818–20 (9th Cir. 2003).

Rather ROSS utterly transformed words into mathematical relationships so that those equations could locate the language of judicial opinions using only the words of judicial opinions. *See* Statement of Facts § L. That kind of change to the works far outpaces what courts have found transformative.

The transformation accomplished by ROSS is analogous to *Sony*, 203 F.3d at 607, and *Sega Enters. v. Accolade, Inc.*, 977 F.2d 1510 (9th Cir. 1992). In *Sega*, the defendant disassembled object code to understand the “ideas and functional concepts embodied in the code.” 977 F.2d at 1520. It then used what it learned from those ideas and concepts to create competing games that could be used on the Sega Genesis game console. *Id.* at 1522-23. The court found fair use because “[w]here there is good reason for studying or examining the unprotected aspects of a copyrighted computer program, disassembly for purposes of such study or examination constitutes a fair use.” *Id.* at 1520. As in *Sega*, ROSS used elements of Plaintiffs’ work for a “legitimate, essentially non-exploitative purpose,” *id.* at 1523, to study the underlying words of judicial opinions—uncopyrightable material. *See id.* at 1522. That strongly favors a finding of fair use.

²¹ Holding modified by *Andy Warhol Found. for the Visual Arts, Inc. v. Goldsmith*, 992 F.3d 99 (2d Cir. 2021), and holding modified by *Andy Warhol Found. for Visual Arts, Inc. v. Goldsmith*, 11 F.4th 26 (2d Cir. 2021).

In *Sony*, the court went even further. There, it found that repeatedly copying Sony’s PlayStation code to understand how it worked was fair use. *Sony*, 203 F.3d at 606-07. It found fair use “notwithstanding the similarity of uses and functions between the Sony PlayStation and the Virtual Game Station,” *id.* at 606, because the Virtual Game Station was a “wholly new product,” and did not contain any of the protectible parts of Sony’s copyrighted code. *Id.* at 606-07. This “use of the copyrighted material was an intermediate one, and thus was only ‘indirect or derivative.’” *Sony*, 203 F.3d at 607 (quoting *Sega*, 977 F.2d at 1522). The court concluded “[w]e are therefore at a loss to see how [plaintiff’s] drafting of entirely new object code for its VGS program could not be transformative, despite the similarities in function and screen output.” *Sony*, 203 F.3d at 606-07. As in *Sony*, ROSS’s product also has, at a general level, similarities in function to Westlaw (both are legal search products). But ROSS’s product is “wholly new” compared to Plaintiffs’ and just like the Virtual Game System, it does not use any of Plaintiffs’ copyrights.²²

²² Insofar as LegalEase used key numbers to identify new cases, that use was not infringement. Neither ROSS nor LegalEase copied the key number system, or any of its component parts, to distribute it to others, or to mimic its organizational scheme to sell a competing organizational scheme to others. Instead, they used the key number system *as Westlaw intended for it to be used*; that is, to identify other cases and to allow a user to click through the system to reach them. See Statement of Facts §§ G-J. Copyright, even when it protects the original elements of a system, does not protect against others *using* that system. *Baker v. Selden*, 101 U.S. 99, 100-01 (1879). “Copyright laws afford protection for a limited time against the publication only and not against the use of a system or plan or idea of which the work is an exposition.” *Taylor v. Commissioner*, 51 F.2d 915, 917 (3d Cir. 1931); *see also Google LLC*, 141 S. Ct. at 1207-08; *Bikram’s Yoga Coll. of India v. Evolation Yoga, Ltd. Liab. Co.*, 803 F.3d 1032, 1040 (9th Cir. 2015); *ACT, Inc. v. Worldwide Interactive Network, Inc.*, 46 F.4th 489, 500 (6th Cir. 2022) (“copyright affords protection only to an author’s *expression* of a system—not the system itself.”) (emphasis in original); *id.* at 505-06 (“Protecting the mere decision to test [a selection] would be akin to granting a patent to ACT on such a testing system—not a copyright protecting original expression.”) (citation omitted); Nimmer on Copyright § 2.18[A] (“Where the owner of a patent obtains the right to exclude others from using the invention, the rights granted to a copyright owner under [§] 106 of the Copyright Act do not include the right to prevent others from using the copyrighted work.”).

As in *Sony* and *Sega*, ROSS’s final product (including its code and the repository of cases that it searches across) contains no headnotes, key numbers, synopses, or any other infringing material. *See* Statement of Facts § L. It also contains none of the key number system’s structure or organization because any such organization was abandoned and discarded through the creation process. *See id.* And ROSS’s final product also does not output any headnotes, key numbers, synopses, or any other infringing material. *See id.* § L.5.

Any copying of protectible material was done for the sake of studying and extracting abstract relationships about the text, not to use the text of judicial opinions and questions. *See* § L. Featurizing the question-and-answer pairs, to measure statistical relationships in judicial language for the purpose of building a machine-learning tool that predicts which judicial passages respond to legal questions, is thus not a copyright *use* of the key numbers or headnotes or other Westlaw Content. Featurizing text does not rely on any of Plaintiffs’ protected expression, and the artificial intelligence that ROSS created does not rely on any of Plaintiffs’ protected expression. *Id.* The final ROSS product contained none of Plaintiffs’ expression, protected or otherwise. *Id.*

Thus, any copying was at most intermediate and this factor favors ROSS.

IV. FACTOR 4: THE EFFECT OF THE USE ON THE MARKET

Under factor four, “[t]he key question is whether defendant’s utilization overlaps in function with plaintiff’s work and thereby supplants either its current markets or the prospective markets to which plaintiff is rightly entitled.” Nimmer on Copyright § 13F.08[E].

If the defendant has not committed this type of harm, which Nimmer calls a “substitutionary harm,” *id.* § 13F.08[A], then this factor favors fair use. *See, e.g., HathiTrust*, 755 F.3d at 100 (factor four favors fair use because “the full-text search function does not serve as a substitute for the books that are being searched”); *Authors Guild*, 804 F.3d at 224-25 (factor

four favors fair use because court did not find that the “snippet view could provide a significant substitute for the purchase of the author's book”).

Lawyers or legal academics that prefer legal research using headnotes and the key number system still have *only* one place to go to use these tools—Westlaw. Statement of Facts § C. Using ROSS’s legal search tool will be of no interest to these legal researchers. *See* Statement of Facts § L. So, the value that Plaintiffs gain from Westlaw Content in the market remains unchanged. *See e.g., Brown v. Netflix, Inc.*, 462 F. Supp. 3d 453, 463 (S.D.N.Y. 2020), *aff’d*, 855 F. App’x 61 (2d Cir. 2021) (“the fourth factor weighs decisively in favor” of fair use where the infringing film “targets a different audience from” plaintiffs’ audience). If any Westlaw users did ultimately leave Westlaw to license ROSS, they did so because ROSS introduced competition for a new and different product into the market. This is virtue under the copyright laws, not a vice. *Sony*, 203 F.3d at 607 (“some economic loss by Sony as a result of this competition does not compel a finding of no fair use. Sony understandably seeks control over the market for devices that play games Sony produces or licenses. The copyright law, however, does not confer such a monopoly”). As in *Sega*, ROSS has built “a legitimate competitor” where nothing in ROSS’s end-product bears expressive similarity to Plaintiffs’ copyrights. *See Sega*, 977 F.2d at 1523.

ROSS’s intermediary use also was not supplanting Plaintiffs’ current market for Westlaw Content. ROSS used the Westlaw Content to develop its training data. *See* Statement of Facts § E-J, L. This includes LegalEase’s use to prepare memos. *See id.* §§ E-J. And it includes ROSS extracting the text from some (though not all) of these memos to generate numerical features, and then running those features through its machine-learning tool. *See id.* § L. Thus, the infringing use was not a *substitute* for Plaintiffs’ use of Westlaw Content in the current market. Plaintiffs do not actually deny this conclusion.

Instead, Plaintiffs state that ROSS supplanted their use of Westlaw Content as AI training data. *See e.g.*, Canter Decl. Ex. 66 (Rebuttal Expert Report of James Malackowski (Sept. 6, 2022)), § 8. However, ROSS did not intend to and did not sell or license any Westlaw Content (or any content) as AI training data and has never had any intention to do so. Ovbiagele Decl. ¶¶ 41-44. ROSS cannot have supplanted Plaintiffs’ position in this prospective market if ROSS never entered it. Cox Decl. ¶¶ 11-13, 14-22; *see also Keck v. Mix Creative Learning Ctr., L.L.C.*, No. 23-20188, --- F.4th ---, 2024 WL 4222188, at *5 (5th Cir. Sept. 18, 2024) (courts should consider whether “widespread use of the work in the same infringing fashion” would affect market for the original); *TCA Television Corp. v. McCollum*, 839 F.3d 168, 186 (2d Cir. 2016) (“a court properly considers the challenged use’s impact on potential licensing revenues for traditional, reasonable, or likely to be developed markets.”) (internal citations and quotation marks omitted).

Moreover, Plaintiffs have no intent of selling the AI training data used by ROSS, and never had any such intent either. When a party expresses an intention to *not* enter a market, the court should accept this decision. *See Cambridge Univ. Press v. Patton*, 769 F.3d 1232, 1277 (11th Cir. 2014) (“if a copyright holder has not made a license available to use a particular work in a particular manner, the inference is that the author or publisher did not think that there would be enough such use to bother making a license available. In such a case, there is little damage to the publisher's market when someone makes use of the work in that way without obtaining a license, and hence the fourth factor should generally weigh in favor of fair use.”); *Swatch Grp. Mgmt. Servs. Ltd. v. Bloomberg L.P.*, 756 F.3d 73, 91 (2d Cir. 2014) (holding that factor four favors fair use “especially in view of the obvious and furthermore conceded fact that Swatch had

no interest in the exploitation of the copyright-protected aspects of the” sound recording at issue).

In any case, AI training data should not be the focus for determining factor four’s application. Courts do not look to an infringer’s use at an intermediary step in the process of the alleged infringement to determine whether the market for or value of the work has been impacted. For example, *Sony* and *Sega* looked at the market for video games, not for software code development. *See e.g., Sega*, 977 F.3d at 1523; *Sony*, 203 F.3d at 607. Plaintiffs’ approach here is just inviting the sort of “circular” reasoning that Nimmer warns courts not to engage in—where the plaintiff defines a potential market as just how the defendant used the work. *See Nimmer on Copyright* § 13F.08[E] (“The inquiry into potential markets gives rise to a danger of circularity—a potential market, no matter how unlikely, has always been supplanted in every fair use case, to the extent that defendant, by definition, has made some actual use of plaintiff’s work, which use could, in turn, be defined in terms of the relevant potential market.”).

Two additional considerations tilt factor four tilts in ROSS’s favor. Factor four “bears a close relationship” to factor one. *Sega*, 977 F.2d at 1523 (factors one and four “accommodate[] the distinction between the copying of works in order to make independent creative expression possible and the simple exploitation of another’s creative efforts”). Thus, ROSS’s transformative use of Westlaw Content closely bears on the analysis here. *Sony*, 203 F.3d at 607 (“because the Virtual Game Station is transformative, and does not merely supplant the PlayStation console, the Virtual Game Station is a legitimate competitor in the market for platforms on which Sony and Sony-licensed games can be played.”); *Sega*, 977 F.2d at 1523.

Finally, the Court also considers here “the public benefits the copying will likely produce.” *Google*, 593 U.S. at 35. Making legal opinions easier to find would increase access to

legal services. *See Georgia v. Public.Resource.Org, Inc.*, 140 S. Ct. 1498, 1507 (2020); Cox Decl. ¶¶ 23-26. Because copyright law is ultimately meant to benefit the public at large, the benefit to the public from increased access to legal services should favor a finding of fair use

CONCLUSION

For the reasons set forth above, ROSS respectfully requests that the Court grant summary judgment in its favor.

OF COUNSEL:

Warrington S. Parker III
Joachim B. Steinberg
Jacob Canter
CROWELL & MORING LLP
3 Embarcadero Ctr., 26th Floor
San Francisco, CA 94111
Tel: (415) 986-2800

Keith J. Harrison
Crinesha B. Berry
CROWELL & MORING LLP
1001 Pennsylvania Avenue NW
Washington, DC 20004
Tel: (202) 624-2500

Dated: October 1, 2024
11790294 / 20516.00001

Public Version Dated: October 9, 2024

Respectfully submitted,

POTTER ANDERSON & CORROON LLP

By: /s/ David E. Moore

David E. Moore (#3983)
Bindu A. Palapura (#5370)
Andrew L. Brown (#6766)
Hercules Plaza, 6th Floor
1313 N. Market Street
Wilmington, DE 19801
Tel: (302) 984-6000
dmoore@potteranderson.com
bpalapura@potteranderson.com
abrown@potteranderson.com

*Attorneys for Defendant/Counterclaimant
ROSS Intelligence, Inc.*